Distributivity facilitates ACD resolution

Jad Wehbe
Massachusetts Institute of Technology
Cambridge, MA, US
jadwehbe@mit.edu

Kate Kinnaird
Massachusetts Institute of Technology
Cambridge, MA, US
kinnaird@mit.edu

Martin Hackl
Massachusetts Institute of Technology
Cambridge, MA, US
hackl@mit.edu

Abstract

Theories of plural predication disagree on the division of labor between the lexical semantics of predicates and various syntactic operators, e.g. the distributivity operator (Krifka 1992; Kratzer 2007). A central issue is whether the LFs giving rise to certain distributive and cumulative readings involve syntactic movement (Beck and Sauerland 2000; Winter 2000). In this project, we adopt the logic of Hackl, Koster-Hale, and Varvoutis (2012) to diagnose movement indirectly via a facilitation effect on ACD resolution. We show that, similar to *every*, definite plurals under a particular distributive construal decrease the cost of ACD resolution compared to a definite singular baseline. This provides evidence that the distributive reading requires LF-movement of the plural DP hosting ACD above the matrix antecedent of the ellided constituent. We argue that this follows naturally only from theories of distributivity/cumulativity on which there is a syntactically represented distributivity operator that the plural DP needs to outscope in order to get the desired reading.

1 Introduction

Plural predication gives rise to cumulative inferences, as illustrated in (1): if (1-a) is true and (1-b) is true, we can infer (1-c). One perspective attributes these inferences to the lexical semantics of predicates. For a 1-place predicate, this amounts to assuming that whenever the predicate is true of parts of a plurality it has to also be true of the resulting plurality. This property can be generalized to n-place predicates as in (2) and has been argued to be a universal property of all lexical predicates (Krifka 1992; Kratzer 2007).[1]

(1) a. Jack invited Student 1 (yesterday).
b. Jack invited Student 2 (today).
c. Jack invited the two students.

(2) Lexical Cumulativity Universal:
$\forall a_1, a'_1, ..a_n, a'_n : \forall P : P(a_1)(..)(a_n) = 1 \wedge P(a'_1)(..)(a'_n) = 1 \Rightarrow P(a_1 + a'_1)(..)(a_n + a'_n) = 1$

To spell out such an account in more concrete terms, we can assume that a sentence like (1-a) has the LF in (3-a). We assume that perfective aspect denotes an existential quantifier over subparts of the time interval that tense picks out and that predicates have a time argument in addition to their individual arguments. The resulting truth-conditions for (1-a) are given in (3-e). The LF for (1-c), even when understood distributively, would be parallel to (3-a) in that the predicate can combine directly with the plural DP, (4-a), and still give rise to cumulative

[1] We illustrate cumulative inference over the object DP in line with our experimental items. However, parallel consideration arise for DPs in other positions including subject DPs.

inferences. I.e. as long as the temporal argument of *invite* falls under (2), (1-c) is true if Student 1 was invited by Jack at some time and Student 2 at some, possibly different time.

(3) **Baseline assumptions:**

a. LF for (1-a): [[Past t][PFV [Jack [invited Student 1]]]]
b. $[\![\text{PFV}]\!] = \lambda P_{\langle i,t \rangle}.\lambda t.\exists t' \subseteq t : P(t')$
c. $[\![\text{Past}]\!]^c = \lambda t : t < t_c.t$
d. $[\![\text{invited}]\!] = \lambda y.\lambda x.\lambda t.x$ invited y at t.
e. $[\![$ (3-a) $]\!]$=1 iff $\exists t' \subseteq t : [\![\text{invited}]\!](s_1)(j)(t')$, where t is a past time interval.

(4) **Lexical cumulativity account:**

a. LF for (1-c): [[Past t] [PFV [Jack [invited the two students]]]]
b. $[\![$ (1-c) $]\!]$=1 iff $\exists t' \subseteq t : [\![\text{invited}]\!](s_1 + s_2)(j)(t')$, where t is a past time interval.
c. $[\![\text{invited}]\!](s_1)(j)(t_1) \wedge [\![\text{invited}]\!](s_2)(j)(t_2) \Rightarrow [\![\text{invited}]\!](s_1 + s_2)(j)(t_1 + t_2)$

A prominent alternative to the lexical perspective is to assume that these inferences are not due to the predicates themselves, but can be attributed to independently needed pluralization operators that modify the predicate (e.g. Beck and Sauerland 2000). On this view, the expression that furnishes cumulative inferences is syntactically complex and we need to ask what the syntactic distribution of the relevant pluralization operators are. One version of the syntactic analysis assumes that pluralization operators apply to every predicate of the right type, be they lexical or syntactically derived. Cumulative inferences for (1-c) may then simply arise from applying a pluralization operator directly to the predicate. A more restrictive perspective, argued for in Wehbe (2023), building on Ionin and Matushansky (2002), is to assume that pluralization operators can only apply to predicates that are derived by movement. Wehbe (2024), moreover, argues that in order to ensure that a sentence like (1-c) is true when Student 1 and Student 2 were invited by Jack at different times, the plural DP and the accompanying pluralization operator have to take scope above Aspect, thus allowing us to distribute over the existential quantifier over times. The resulting movement-based analysis is given in (5).

(5) **Movement analysis:**

a. LF: [[Past t] [the two students [* [1 [PFV [Jack [invited t_1]]]]]]]
b. $[\![$ (5-a) $]\!]$=1 iff $\exists t' \subseteq t : [\![\text{invited}]\!](s_1)(j)(t') = 1 \wedge \exists t'' \subseteq t : [\![\text{invite}]\!](s_2)(j)(t'') = 1$
c. $[\![*]\!] = \lambda f_{\langle e,t \rangle}.\lambda x.\exists g_{\langle e,t \rangle}[[\forall y[g(y) = 1 \rightarrow f(y) = 1] \wedge x = \bigoplus g]$

We see that both approaches succeed in deriving cumulative inferences for cases like (1-c). However, they differ with respect to what role the syntax plays. On the lexical cumulativity view, the syntactic structure of the sentence that gives rise to the cumulative inference is not different from the one that is needed for singular object DPs or, for that matter, the one that gives rise to a collective reading on which there are no inferences to the parts of the object plurality. In all three cases, the object DP can be interpreted in situ. For the syntactic approach sketched above, however, deriving the distributive reading involves a logical form in which the object DP is not interpreted in its base position. Rather, on this account, it has to take scope above Aspect to support inferences to its parts.

Since both approaches are equally successful at the basic task of deriving distributive truth-conditions for sentences like (1-c) we need to look elsewhere for evidence that might help identify the correct logical form of such sentences under this interpretation. One fruitful avenue is to look for correlations between the presence/absence of DP movement and the presence/absence of cumulative inferences. Wehbe (2023), for instance, observes that in the Lebanese Arabic double-subject construction cumulative inferences are present for plural DPs that have moved but absent for those that are interpreted in situ. Another approach, pursued in the present

study, is to ask whether sentences that are interpreted distributively and thus exhibit cumulative inferences have processing signatures that are indicative of DP movement.

2 Probing for the correct LF via ACD resolution

The present study follows the approach of Hackl, Koster-Hale, and Varvoutis (2012) to look for a processing signature of covert DP-movement. The basic idea in Hackl et al. is to use the relative ease with which comprehenders are able to resolve antecedent contained deletion (ACD) as a proxy for how and when they can determine the correct LF-position of the DP hosting the ACD site. To appreciate their logic, consider the factorial template in (6) in which the presence or absence of ACD is crossed with two types of determiners on the hosting DP - the quantificational determiner *every* or the non-quantificational determiner *the*.

(6) Jane read {every$_{\textbf{+quant.}}$, the$_{\textbf{-quant.}}$} book that Mary { did$_{\textbf{+ACD}}$, wrote$_{\textbf{-ACD}}$}.

For a sentence to exhibit ACD it has to have a structural description (LF) where the elided VP is not included inside the matrix VP, which serves as its antecedent (Sag 1976). This can be achieved, on canonical assumptions, if the object DP hosting the ACD site is moved to a position higher than the matrix VP, thereby "undoing" antecedent containment (Sag 1976, Williams 1977; see Jacobson 1992 for an alternative). Importantly, while both quantificational and non-quantificational DPs need to be moved when they host an ACD site, Hackl et al. assume that only the former require it on their own, irrespective of whether or not they host an ACD site. From a procedural perspective, this means that the [+quant, +ACD] variant of (6) contains two triggers for the parser to identify the correct LF while the [-quant,+ACD] variant contains only one, namely the ACD site itself.

Adopting a linking hypothesis that connects the number of cues for a particular LF with the ease with which that LF can be identified leads Hackl et al. to expect a facilitation effect for *every* on ACD resolution. More specifically, since resolving ACD is a complex process, which involves postulating an LF without antecedent containment (i.e. by assumption covert movement of the object DP), identifying the antecedent VP and retrieving it from memory when computing the meaning of the relative clause, the [+ACD] variants of (6) are expected to be more costly than their [-ACD] counterparts. It is also expected, however, that when the host DP is projected from *every* this cost is reduced and ACD resolution is facilitated since a critical component of resolving ACD - identifying an LF without containment - is triggered independently from and earlier than the ACD site. By contrast, projecting the host DP from *the* does not provide a separate trigger for moving the object DP above the matrix VP. Thus, in this condition the cue for such an LF is provide by the ACD site alone, at a stage when a parse without movement is likely entrenched and so relatively costly to abandon. Combining these two types of considerations leads to the prediction of a particular type of interaction such that the [+quant,+ACD] condition is relatively less costly than its [-quant,+ACD] counterpart. This facilitation effect has been documented, among other things and most relevantly for our purpose here, in naturalness ratings in Hackl et al. and it has been replicated in a series of subsequent work including Gibson et al. (2015), who offer an alternative account of the facilitation effect.[2]

The present project extends this logic to distributive readings of plural definites. Concretely, if distributive readings are due to a syntactically represented pluralization operator that applies to a derived predicate, then a definite plural object DP, just like *every* in (6), triggers movement

[2]The facilitation effect has also been documented in self-paced reading data in Hackl et al. and both types of observations have been replicated and extended to a variety of other expressions that are hypothesized to trigger upstream movement in their own right (see Hackl, Koster-Moeller, and Gottstein 2009, Breakstone et al. 2012, Kotek 2019).

in order to get the desired reading. Thus, we expect a similar facilitation effect to Hackl et al.'s observation for *every* for distributively interpreted definite plurals.

3 Distributivity and ACD resolution: an experimental study

3.1 Task and Design

We adopt the logic of Hackl et al. to probe whether definite plural DPs in object position undergo movement when they are interpreted distributively. We focus on a particular instance of a distributive reading where the different atomic parts of the relevant plural DP participated in the event in question at different times. We will refer to this as a *one-at-a-time* reading. Our target sentences, which parallel the format in (6), are paired with contexts that strongly encourage a *one-at-a-time* interpretation. An example is given in (7), where we crossed +/-ACD and the type of DP (definite plural vs. singular).

(7) **Context:** 25 international students are invited for a campus visit at an Ivy league university, and they get the chance to speak with professors one-on-one.
Sentence: The English literature professor met with the humanities {students$_{\mathbf{+PL}}$, student$_{\mathbf{-PL}}$} that the philosophy professor {did$_{\mathbf{+ACD}}$, interviewed $_{\mathbf{-ACD}}$} due to the dean's insistence.

Items were presented in PCIbex (Zehr and Schwarz 2023) using a counterbalanced Latin Square design, with each participant being assigned to one of four lists. The experiment consisted of 32 fillers and 32 target items. Item order was randomized for every participant. Filler items were of similar complexity and length to target items. As exemplified in (7), Participants first saw a context sentence followed by a target sentence, and then were prompted to rate the target sentence's naturalness relative to the context on a 7 point Likert scale. Following their selection, the context sentence and target sentence were removed from the screen. Participants were then given a Yes/No comprehension question asking about information conveyed by the target sentence. To probe for whether participants are paying attention to the crucial distinction between plural and singular definite DPs, some of our comprehension questions are sensitive to this distinction (e.g. for (7): *Did the English literature professor meet with only one student?*). To establish a baseline for our design (seeking to replicate Hackl et al.), we also ran a variant of the study where the definite plural in (7) is replaced with *every* (e.g. *every humanities student*), keeping everything else constant.

3.2 Results

For each of our two experiments, we recruited 80 monolingual English speakers using crowdsourcing platform Prolific. Participants were excluded if they failed any of our three practice items or if they answered less than 8 out of our 12 attention check items correctly (36 for the baseline and 19 for the target experiment). Participants were paid approximately 5 USD.

Mean ratings for the baseline and target experiments, with 95 percent confidence interval-based error bars, are displayed in Figure 1. We report the maximally-specified linear mixed-effects model (conditioned on convergence), predicting naturalness ratings as a function of [+/-ACD], DP type (every vs. definite singular for the baseline experiment and definite plural vs. definite singular for the target experiment), and their interaction. We include random slopes of DP type and [+/-ACD] for participants and a random slope of [+/-ACD] for items. The model reveals a significant interaction between [+/-ACD] and DP type for both the baseline experiment and the target experiment ($p<0.0001$ and $p<0.05$, respectively). Analyses were performed in R Statistical Software (v 4.2.2; R Core Team 2022) using the lme4 package (Bates et al. 2015).

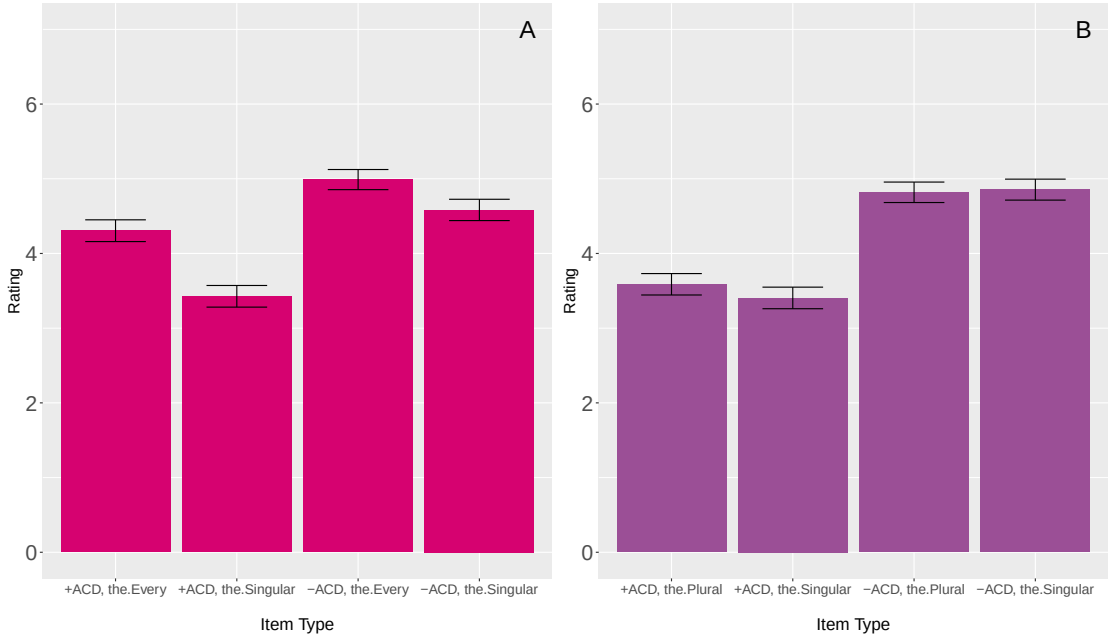

Figure 1: Mean naturalness ratings for Baseline Experiment in A (N=80) and Target Experiment in B (N=80)

3.3 Discussion

Our baseline experiment with *every* replicates the results in Hackl et al. In particular, we see that the presence of ACD results in a marked degradation of naturalness, but critically, we observe an interaction indicating that ACD resolution is facilitated with *every*, when compared to the definite singular. Following Hackl et al., we take this to support the conclusion that *every* in object position has to move for reasons independent of ACD.

Turning to the target experiment, we see a parallel facilitation effect of the definite plural on ACD resolution. Adopting, again, the reasoning in Hackl et al., we take this to show that the distributive reading over a plural object DP in our environment requires movement of that DP. Interestingly, as is evident in Figure 1, the size of the interaction effect with the definite plural is smaller than that with *every* (Model estimate= 0.2188 for the target experiment and 0.47188 for the baseline experiment). This difference in effect size could be due to several factors. One possibility is that that the distinction between *the students* and *the student* is easier to miss for our participants than the distinction between *every student* and *the student*. This would result in more noise in the target experiment, thereby lowering the size of the effect. To explore this possibility, we excluded participants who failed to correctly answer more than one-third of comprehension questions which were sensitive to the singular/plural distinction. This resulted in the effect size going up to 0.26906 and the p-value going down to 0.0194 from 0.0457.[3] This is as expected on the view that part of the difference in effect size is due to participants missing the singular/plural distinction in the target experiment.

Another possible contributing factor to the difference in effect size is that the information needed to recognize the trigger for movement in the two experiments is different. In the case of *every*, it is simply recognizing that *every* occurs in object position. On the other hand, the relevant information in the case of the definite plural crucially includes the distributive context (assuming that definite plurals have to move only under the distributive construal). Facilitation is therefore only predicted for those participants who keep the context in mind and attempt to

[3]Excluding participants at 25% and 50% thresholds resulted in similar changes in effect size and p value, although simpler models were required for convergence.

generate a matching distributive reading for the target sentence.

4 Conclusion

Just like Hackl et al., we observe a link between the nature of the DP hosting an ACD site (definite singular vs definite plural) and the ease of ACD resolution. Specifically, definite plurals under our *one-at-a-time* contexts exhibit a facilitation effect. We take this to show that the definite plural is perceived by comprehenders as a trigger for movement in our environment, allowing the parse where antecedent containment has been undone to be identified before the ACD site is reached. This follows from theories where pluralization operators are always syntactically represented and where movement of the object DP to a position above the relevant operator is required to generate the desired distributive reading (Beck and Sauerland 2000; Wehbe 2023; Wehbe 2024). On the other hand, it is difficult to explain this facilitation effect if our distributive reading is due to lexical cumulativity. It is nevertheless possible that the lexical cumulativity universal holds and that it is the particular type of distributive reading over times that we target here that requires movement. Such a mixed view, however, raises doubts about the explanatory utility of lexical cumulativity. In particular, giving up an account of our cases in terms of lexical cumulativity can only be achieved at the cost of a stipulation, e.g. that temporal specifications of predicates are not covered by lexical cumulativity.

Our results raise interesting questions in the context of previous processing studies of distributivity. For example, Dotlačil and Brasoveanu (2021) argue for a distinction between lexical and phrasal distributivity based on differences in how easy the corresponding readings are to access. In particular, they show that the distributive reading over an indefinite exemplified in (8-a) is dis-preferred relative to the collective reading, while there is no parallel difficulty for the distributive reading in (8-b). Dotlačil and Brasoveanu (2021) take this as evidence that these two kinds of distributive readings are due to distinct sources: the reading in (8-a) requires a distributivity operator to apply to a derived predicate that includes the indefinite, while the distributive reading in (8-b) follows from lexical cumulativity.

(8) a. The students read a book./ Dist. reading: The students each read a book.
b. The students lifted the piano./Dist. reading: The students each lifted the piano.

Given that our diagnostic indicates that the *one-at-a-time* reading is due to phrasal distributivity, if the characterization of Dotlačil and Brasoveanu (2021) above is correct, we would expect our sentences to pattern with (8-a) rather than (8-b). It is not clearthat the *one-at-a-time* distributive reading is in fact dis-preferred in the same way as (8-a). If it turns out that our sentences pattern with (8-b), this would require a different explanation for the processing difference in (8) which nevertheless maintains that both LFs involve a distributivity operator.

To conclude, our study lays the foundation for a new tool to compare frameworks where plural predication is mediated by syntactic operators (e.g. Beck and Sauerland 2000), in which movement plays a crucial role, to frameworks where we don't expect to see systematic syntactic correlates of plural predication (e.g. Winter 2000; Schmitt 2019). Looking ahead, we hope to use this new tool to investigate the role of movement in generating different readings in plural predication. For example, non-local cumulative readings have been argued to require movement of the relevant DPs to a derived scope position (Beck and Sauerland 2000), but certain accounts, including the more recent plural projection framework (Schmitt 2019) are able to generate these non-local readings without movement. Our tool can help to distinguish the predictions of these two theories. Another direction for future research is to investigate whether distributivity is in fact crucial for the facilitation effect with definite plurals. We therefore plan to test whether plural DPs under collective construals give rise to the same facilitation effect we observed here.

References

Bates, Douglas et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Beck, Sigrid and Uli Sauerland (2000). "Cumulation is needed: A reply to Winter (2000)". In: *Natural language semantics* 8.4, pp. 349–371.

Breakstone, Micha Yochanan et al. (2012). "On the Analysis of Scope Ambiguities in Comparative Constructions: Converging Evidence from Real-Time Sentence Processing and Offline Data". In: *Proceedings of SALT*. Vol. 21, pp. 712–731.

Dotlačil, Jakub and Adrian Brasoveanu (2021). "The representation and processing of distributivity and collectivity: ambiguity vs. underspecification". In: *Glossa: A journal of general linguistics* 6.1.

Gibson, Edward et al. (2015). "A pragmatic account of complexity in definite antecedent-contained-deletion relative clauses". In: *Journal of Semantics* 32.4, pp. 579–618.

Hackl, Martin, Jorie Koster-Hale, and Jason Varvoutis (2012). "Quantification and ACD: Evidence from real-time sentence processing". In: *Journal of Semantics* 29.2, pp. 145–206.

Hackl, Martin, Jorie Koster-Moeller, and Andrea Gottstein (2009). "Processing Opacity". In: *Proceedings of Sinn und Bedeutung 13*. Vol. 13. 1, pp. 171–186. DOI: https://doi.org/10.18148/sub/2009.v13i1.524.

Ionin, Tania and Ora Matushansky (2002). "DPs with a twist: A unified analysis of Russian comitatives". In: *Proceedings of FASL*. Vol. 11.

Jacobson, Pauline (1992). "Antecedent Contained Deletion in a Variable-Free Semantics". In: *Proceedings of SALT 2*, pp. 193–213. DOI: https://doi.org/10.3765/salt.v2i0.3027.

Kotek, Hadas (Feb. 2019). *Composing Questions*. The MIT Press. ISBN: 9780262351065. DOI: 10.7551/mitpress/10774.001.0001. URL: https://doi.org/10.7551/mitpress/10774.001.0001.

Kratzer, Angelika (2007). "On the plurality of verbs". In: *Event structures in linguistic form and interpretation* 269.300, p. 2.

Krifka, Manfred (1992). "Thematic Relations as Links between Nominal Reference and Temporal Constitution/Manfred Krifka". In: *Lexical matters* 24, p. 29.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Sag, Ivan A (1976). "Deletion and logical form." PhD thesis. Massachusetts Institute of Technology.

Schmitt, Viola (2019). "Pluralities across categories and plural projection". In: *Semantics and Pragmatics* 12, pp. 17–1.

Wehbe, Jad (2023). "Against the lexical view of cumulative inferences". In: *Proceedings of Sinn und Bedeutung*. Vol. 27, pp. 693–711.

— (2024). "Covert Reciprocals: A scope-based account of reciprocal alternations". In.

Williams, Edwin S (1977). "Discourse and logical form". In: *Linguistic inquiry*, pp. 101–139.

Winter, Yoad (2000). "Distributivity and dependency". In: *Natural language semantics* 8.1, pp. 27–69.

Zehr, Jérémy and Florian Schwarz (Jan. 2023). *PennController for Internet Based Experiments (IBEX)*. DOI: 10.17605/OSF.IO/MD832. URL: osf.io/md832.